

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

**APLICAÇÃO DE MCMC NA ESTIMAÇÃO DE
MÁXIMA VEROSSIMILHANÇA PARA
PROCESSOS AR(p) e MA(q)**

**MARINHO GOMES DE ANDRADE FILHO
MARCIA FUMI MIZOI**

Nº 52

**NOTAS DO ICM C
Série Estatística**

**São Carlos
Out./1998**

Aplicação de MCMC na Estimação de Máxima Verossimilhança para Processos AR(p) e MA(q)

Marinho Gomes de Andrade Filho; Marcia Fumi Mizoi

SCE - ICMC - USP
Cx.P.: 668; 13560-970 São Carlos - SP

Resumo

Neste trabalho, abordamos os modelos de séries temporais estacionárias do tipo AR(p) e MA(q). O objetivo principal é propor para estes modelos o uso de algoritmos de simulação de Monte Carlo com Cadeia de Markov (MCMC) para o cálculo das estimativas de máxima verossimilhança. Aqui, os algoritmos utilizados foram o amostrador de Gibbs e o algoritmo de Metropolis-Hastings. Os resultados obtidos usando MCMC são comparados com as estimativas feitas por métodos numéricos propostos em [Box, Jenkins e Reinsel - 1994] e [Miller - 1995].

Abstract

In this work, the autoregressive and moving average time series models were considered. The main objective here is to use the Markov Chain Monte Carlo (MCMC) method (Metropolis-Hastings algorithm and the Gibbs Sampler) to calculate the maximum likelihood estimates in the ordinary sense (a point in the parameter space that maximizes the likelihood). The proposed method was applied to the simulated and the real date record. The estimates obtained by the MCMC method were compared with likelihood estimates by iterative numerical methods.

Palavras-chaves: estimativa de máxima verossimilhança; MCMC; modelos de séries temporais.

1. Introdução

Na análise de séries temporais, um dos passos importantes é a estimação dos parâmetros do modelo matemático ajustado à série dada. Uma maneira de proceder esta

estimação é através do uso de métodos de estimativas de máxima verossimilhança. No nosso caso estamos interessados em obter as estimativas para os modelos de séries temporais estacionárias do tipo AR(p) e MA(q). A complexidade da expressão da função de verossimilhança exata para estes modelos sugere o uso de métodos numéricos baseados em aproximações como os apresentados em [Box, Jenkins e Reinsel - 1994]. Também em [Miller - 1995] é proposto um método numérico, para o caso de modelos AR(p), que utiliza um algoritmo iterativo baseado em expressões mais simples para as derivadas da função de verossimilhança. Um método proposto para estimar os parâmetros de modelos MA(1) é apresentado em [Anderson, et al - 1996]. O objetivo deste projeto é propor como método de obtenção das estimativas de máxima verossimilhança, o uso de algoritmos de simulação de Monte Carlo com Cadeia de Markov (MCMC), mais especificamente o algoritmo amostrador de Gibbs [Casella e George - 1992] e o algoritmo de Metropolis-Hastings [Chib e Greenberg - 1995]. O uso de simulação de Monte Carlo para o cálculo de estimativas de máxima verossimilhança foi proposto inicialmente por [Geyer e Thompson- 1992] e [Geyer - 1994]. Nesse trabalho é usado programação linear para encontrar o suporte da função de verossimilhança e então acha a distribuição da família exponencial neste suporte. Na proposta do nosso trabalho os parâmetros dos modelos são interpretados como variáveis aleatórias cuja densidade tem como núcleo a função de verossimilhança do modelo, então os algoritmos amostrador de Gibbs e Metropolis-Hastings são usados para gerar uma amostra desses parâmetros e a estimativa de máxima verossimilhança desses parâmetros é tomada como sendo a moda da amostra gerada.

No final do trabalho, apresentamos comparações entre os resultados da aplicação dos métodos numéricos e da aplicação de MCMC em séries geradas por modelos AR(2) e MA(2).

2. A Função de Verossimilhança Exata

2.1 Função de Verossimilhança Exata para Processos AR(p)

Suponhamos que uma dada série $\mathbf{z}_n = (z_1, z_2, \dots, z_n)'$ é gerada pelo modelo autorregressivo estacionário de ordem p, ou AR(p):

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t$$

onde $E[z_t] = 0$ e $\{a_t\}$ é um ruído branco gaussiano com variância σ_a^2 . Em [Box, Jenkins e Reinsel - 1994] mostra-se que a função de verossimilhança exata para este modelo pode ser expressa na forma

$$L(\boldsymbol{\phi}, \tau | \mathbf{z}_n) = \left(\frac{\tau}{2\pi}\right)^{n/2} \left| \mathbf{M}_p^{(p)} \right|^{1/2} \exp \left\{ -\frac{\tau}{2} \left[\mathbf{z}'_p \mathbf{M}_p^{(p)} \mathbf{z}_p + \sum_{t=p+1}^n (z_t - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p})^2 \right] \right\} \quad (1)$$

onde $\boldsymbol{\phi} = (\phi_1, \phi_2, \dots, \phi_p)'$, $\tau = \frac{1}{\sigma_a^2}$ e a matriz $\mathbf{M}_p^{(p)}$ é dada por:

$$\mathbf{M}_p^{(p)} = \{\gamma_{|i-j|}\}^{-1} \sigma_a^2 = \begin{bmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{p-2} \\ \vdots & \vdots & \dots & \vdots \\ \gamma_{p-1} & \gamma_{p-2} & \dots & \gamma_0 \end{bmatrix}^{-1} \sigma_a^2$$

com $\gamma_0, \gamma_1, \dots, \gamma_{p-1}$ as autocovariâncias do processo.

2.2 Função de Verossimilhança Exata para Processos MA(q)

Consideremos o modelo de médias-móveis invertível de ordem q, MA(q), dado por:

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

onde $E[z_t] = 0$ e suponhamos que uma dada série $\mathbf{z}_n = (z_1, z_2, \dots, z_n)'$ é gerada por este modelo. Suponhamos também que $\{a_t\}$ é um ruído branco gaussiano com variância σ_a^2 . Seja $\mathbf{a}_* = (a_{1-q}, a_{2-q}, \dots, a_0)'$ um vetor q-dimensional dos valores preliminares do ruído. Então, pode-se mostrar que a densidade conjunta de \mathbf{z}_n e \mathbf{a}_* pode ser escrita como:

$$p(\mathbf{z}_n, \mathbf{a}_* | \boldsymbol{\theta}, \tau) = \left(\frac{\tau}{2\pi}\right)^{(n+q)/2} \exp \left\{ -\frac{\tau}{2} [(\mathbf{z}_n - \mathbf{F}\mathbf{a}_*)' \mathbf{G}'^{-1} \mathbf{G}^{-1} (\mathbf{z}_n - \mathbf{F}\mathbf{a}_*) + \mathbf{a}_*' \mathbf{a}_*] \right\}$$

onde:

- $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_q)'$;
- \mathbf{G} é uma matriz triangular inferior de tamanho $n \times n$ formada por 1's na diagonal principal, $-\theta_1$ na primeira subdiagonal, $-\theta_2$ na segunda subdiagonal, e assim por diante, com $\theta_i = 0$, para $i > q$;
- \mathbf{F} é uma matriz de tamanho $n \times q$ da forma $\mathbf{F} = (\mathbf{B}'_q, \mathbf{O}')'$, onde

$$\mathbf{B}_q = \begin{bmatrix} \theta_q & \theta_{q-1} & \dots & \theta_1 \\ 0 & \theta_q & \dots & \theta_2 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \theta_q \end{bmatrix}$$

é uma matriz quadrada de ordem q e O uma matriz nula de ordem $(n - q) \times q$.

Esta densidade pode ser escrita na seguinte forma:

$$p(\mathbf{z}_n, \mathbf{a}_* | \boldsymbol{\theta}, \tau) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} [(\mathbf{z}_n - \mathbf{F}\mathbf{a}_*)' \mathbf{G}'^{-1} \mathbf{G}^{-1} (\mathbf{z}_n - \mathbf{F}\mathbf{a}_*)]\right\} \\ \cdot \left(\frac{\tau}{2\pi}\right)^{q/2} \exp\left\{-\frac{\tau}{2} \mathbf{a}_*' \mathbf{a}_*\right\}$$

a qual pode ser interpretada como

$$p(\mathbf{z}_n, \mathbf{a}_* | \boldsymbol{\theta}, \tau) = p(\mathbf{z}_n | \mathbf{a}_*, \boldsymbol{\theta}, \tau) p(\mathbf{a}_* | \boldsymbol{\theta}, \tau)$$

onde:

$$p(\mathbf{z}_n | \mathbf{a}_*, \boldsymbol{\theta}, \tau) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} [(\mathbf{z}_n - \mathbf{F}\mathbf{a}_*)' \mathbf{G}'^{-1} \mathbf{G}^{-1} (\mathbf{z}_n - \mathbf{F}\mathbf{a}_*)]\right\}$$

e

$$p(\mathbf{a}_* | \boldsymbol{\theta}, \tau) = \left(\frac{\tau}{2\pi}\right)^{q/2} \exp\left\{-\frac{\tau}{2} \mathbf{a}_*' \mathbf{a}_*\right\} \quad (2)$$

Supondo conhecido o valor de \mathbf{a}_* , podemos estimar os parâmetros $\boldsymbol{\theta}$ e τ usando a função de verossimilhança exata:

$$L(\boldsymbol{\theta}, \tau | \mathbf{z}_n, \mathbf{a}_*) = p(\mathbf{z}_n | \mathbf{a}_*, \boldsymbol{\theta}, \tau) p(\mathbf{a}_* | \boldsymbol{\theta}, \tau)$$

ou seja,

$$L(\boldsymbol{\theta}, \tau | \mathbf{z}_n, \mathbf{a}_*) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} [(\mathbf{z}_n - \mathbf{F}\mathbf{a}_*)' \mathbf{G}'^{-1} \mathbf{G}^{-1} (\mathbf{z}_n - \mathbf{F}\mathbf{a}_*)]\right\} \\ \cdot \left(\frac{\tau}{2\pi}\right)^{q/2} \exp\left\{-\frac{\tau}{2} \mathbf{a}_*' \mathbf{a}_*\right\} \quad (3)$$

3. Método de Simulação de Monte Carlo com Cadeia de Markov (MCMC) para Estimativas de Máxima Verossimilhança

Consideremos a função de verossimilhança $L(\boldsymbol{\lambda})$, $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$ (espaço de parâmetros), para os modelos AR(p) ou MA(q). Interpretando os parâmetros destes modelos como variáveis aleatórias, podemos construir uma função densidade de probabilidade por:

$$p(\lambda) = \frac{1}{c} L(\lambda)$$

onde

$$c = \int_{\lambda} L(\lambda) d\lambda$$

Desta forma, o estimador de máxima verossimilhança $\hat{\lambda}$ é a moda da densidade $p(\lambda)$.

A vantagem desta interpretação é que podemos calcular $\hat{\lambda}$ usando algoritmos de simulação MCMC, sem necessidade do conhecimento da constante c . Esses algoritmos

simulam no espaço de parâmetros uma cadeia de Markov cuja distribuição estacionária seja dada por p , de modo que, para uma iteração suficientemente alta da cadeia podemos considerar que os valores finais simulados representam amostras da densidade desejada. Os algoritmos aqui utilizados foram o amostrador de Gibbs [Casella e George - 1992] e o algoritmo de Metropolis-Hastings [Chib e Greenberg - 1995]. Para identificar e monitorar a convergência destes algoritmos, utilizamos o critério de convergência proposto por [Gelman e Rubin -1992]. Este método propõe o cálculo do fator de redução de escala $\sqrt{\hat{R}}$, cujo valor deve estar próximo de 1 para garantir a convergência. A seguir descrevemos o desenvolvimento da teoria necessária para a aplicação dos algoritmos nos modelos AR(p) e MA(q).

3.1 Aplicação de MCMC em Processos AR(p)

Consideremos a expressão da função de verossimilhança exata apresentada em (1). As estimativas de máxima verossimilhança podem ser obtidas, aplicando-se o algoritmo de Metropolis-Hastings. Inicialmente, vamos reescrever a função de verossimilhança como o produto de duas funções. Para isso denotemos:

$$Y = \begin{bmatrix} z_{p+1} \\ z_{p+2} \\ \vdots \\ z_n \end{bmatrix}_{(n-p) \times 1}, \quad X = \begin{bmatrix} z_p & \dots & z_1 \\ z_{p+1} & \dots & z_2 \\ \vdots & \dots & \vdots \\ z_{n-1} & \dots & z_{n-p} \end{bmatrix}_{(n-p) \times p} \quad e \quad \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix}_{p \times 1}$$

Podemos então, escrever o somatório da expressão na forma

$$\sum_{t=p+1}^n (z_t - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p})^2 = (Y - X\phi)'(Y - X\phi)$$

É possível mostrar que:

$$(Y - X\phi)'(Y - X\phi) = (\phi - \hat{\phi})' X' X (\phi - \hat{\phi}) + (Y - \hat{Y})'(Y - \hat{Y})$$

onde

$$\hat{\phi} = (X'X)^{-1}X'Y$$

e

$$\hat{Y} = X\hat{\phi} = X(X'X)^{-1}X'Y$$

O vetor $\hat{\phi}$ é denominado o estimador de máxima verossimilhança quando usamos a função de verossimilhança "aproximada" - [Box, Jenkins e Reinsel - 1994].

Assim, podemos escrever a função de verossimilhança exata como:

$$L(\phi, \tau | \mathbf{z}_n) = \Psi_1(\phi, \tau) \Psi_2(\phi, \tau)$$

definindo-se as funções

$$\Psi_1(\phi, \tau) = \left(\frac{\tau}{2\pi}\right)^{(n-1)/2} |M_p^{(p)}|^{1/2} \exp\left\{-\frac{\tau}{2} [\mathbf{z}'_p M_p^{(p)} \mathbf{z}_p + (Y - \hat{Y})'(Y - \hat{Y})]\right\}$$

e

$$\Psi_2(\phi, \tau) = \left(\frac{\tau}{2\pi}\right)^{1/2} \exp\left[-\frac{\tau}{2}(\phi - \hat{\phi})'X'X(\phi - \hat{\phi})\right]$$

Notemos que $\Psi_2(\phi, \tau)$ é uma normal p-dimensional com média $\hat{\phi}$ e precisão $\tau X'X$.

Seja $p(\phi, \tau) = \Psi_2(\phi, \tau)$. Temos que

$$p(\phi|\tau) \sim N(\hat{\phi}, (\tau X'X)^{-1})$$

$$p(\tau|\phi) \sim \Gamma\left(\frac{3}{2}, \frac{1}{2}(\phi - \hat{\phi})'X'X(\phi - \hat{\phi})\right)$$

Então, podemos aplicar o algoritmo de Metropolis-Hastings utilizando as densidades acima para gerar os candidatos da amostra, sendo que a probabilidade de aceitação de cada candidato $\phi^{(i)} \sim p(\phi|\tau)$ é dada por:

$$\alpha_1 = \min\left\{1, \frac{\Psi_1(\phi^{(i)}, \tau)}{\Psi_1(\phi^{(i-1)}, \tau)}\right\}$$

e a probabilidade de aceitação de $\tau^{(i)} \sim p(\tau|\phi)$ é dada por:

$$\alpha_2 = \min\left\{1, \frac{\Psi_1(\phi, \tau^{(i)})}{\Psi_1(\phi, \tau^{(i-1)})}\right\}$$

3.2 Aplicação de MCMC em Processos MA(q)

Vimos anteriormente em (3) a expressão da função de verossimilhança exata para um processo MA(q). Observamos que:

$$p(\boldsymbol{\theta} | \tau, \mathbf{a}_*, \mathbf{z}_n) \propto \tau^{n/2} \exp\left\{-\frac{\tau}{2} [(\mathbf{z}_n - F\mathbf{a}_*)' \mathbf{G}'^{-1} \mathbf{G}^{-1} (\mathbf{z}_n - F\mathbf{a}_*)]\right\}$$

e

$$p(\tau | \boldsymbol{\theta}, \mathbf{a}_*, \mathbf{z}_n) \propto \tau^{(n+q)/2} \exp\left\{-\frac{\tau}{2} S(\boldsymbol{\theta}, \mathbf{a}_*)\right\}$$

onde

$$S(\boldsymbol{\theta}, \mathbf{a}_*) = (\mathbf{z}_n - F\mathbf{a}_*)' \mathbf{G}'^{-1} \mathbf{G}^{-1} (\mathbf{z}_n - F\mathbf{a}_*) + \mathbf{a}_*' \mathbf{a}_*$$

Também em (2), vimos que

$$p(\mathbf{a}_* | \boldsymbol{\theta}, \tau) = \left(\frac{\tau}{2\pi}\right)^{q/2} \exp\left\{-\frac{\tau}{2} \mathbf{a}_*' \mathbf{a}_*\right\}$$

Então, utilizando algoritmos MCMC, podemos gerar valores para \mathbf{a}_* , τ e $\boldsymbol{\theta}$, implementando o seguinte algoritmo:

1) Fazer $j = 1$ e dar os valores iniciais $\mathbf{a}_*^{(0)}$, $\tau^{(0)}$ e $\boldsymbol{\theta}^{(0)}$;

2) Gerar:

$$\mathbf{a}_*^{(j)} \sim p\left(\mathbf{a}_* | \boldsymbol{\theta}^{(j-1)}, \tau^{(j-1)}\right);$$

$$\tau^{(j)} \sim p\left(\tau | \boldsymbol{\theta}^{(j-1)}, \mathbf{a}_*^{(j)}, \mathbf{z}_n\right);$$

$$\boldsymbol{\theta}^{(j)} \sim p\left(\boldsymbol{\theta} | \tau^{(j)}, \mathbf{a}_*^{(j)}, \mathbf{z}_n\right);$$

3) Fazer $j \leftarrow j + 1$ e repetir os passos 2) e 3) até convergência.

Note que:

$$\mathbf{a}_* \sim N(0, \tau^{-1})$$

$$\tau \sim \Gamma\left(\frac{n+q}{2} + 1, \frac{S(\boldsymbol{\theta}, \mathbf{a}_*)}{2}\right)$$

$$\boldsymbol{\theta} \sim p(\boldsymbol{\theta} | \tau, \mathbf{z}_n, \mathbf{a}_*).$$

Então, podemos utilizar o amostrador de Gibbs para gerar \mathbf{a}_* e τ e, no caso de $\boldsymbol{\theta}$, como não é possível identificar uma distribuição conhecida, utilizamos o algoritmo de Metropolis-Hastings tomando como núcleo uma distribuição normal $N(\mu_\theta, \sigma_\theta^2)$ onde os parâmetros μ_θ e σ_θ^2 devem ser ajustados de acordo com a taxa de aceitação dos valores simulados.

4. Exemplos

Apresentamos aqui a aplicação do método MCMC para o caso específico de séries temporais geradas por modelos AR(2) e MA(2). Para o modelo AR(2), os resultados deste método são comparados com os resultados obtidos pelos métodos numéricos estimativas mínimos quadrados e aproximação das estimativas de máxima verossimilhança apresentados em [Box, Jenkins e Reinsel - 1994], e pelo método ARMLE apresentado em [Miller - 1995]. Para o modelo MA(2), os métodos numéricos utilizados para comparação foram: estimação de máxima verossimilhança condicional e estimação de máxima verossimilhança incondicional também propostos em [Box, Jenkins e Reinsel - 1994].

Exemplo 1:

Modelo AR(2): série de dados reais com 576 observações formada pelas vazões médias mensais do reservatório de Furnas durante 48 anos, no período de janeiro de 1931 a dezembro de 1978. A seguir, apresentamos o gráfico da série e os resultados de cada método aplicado.

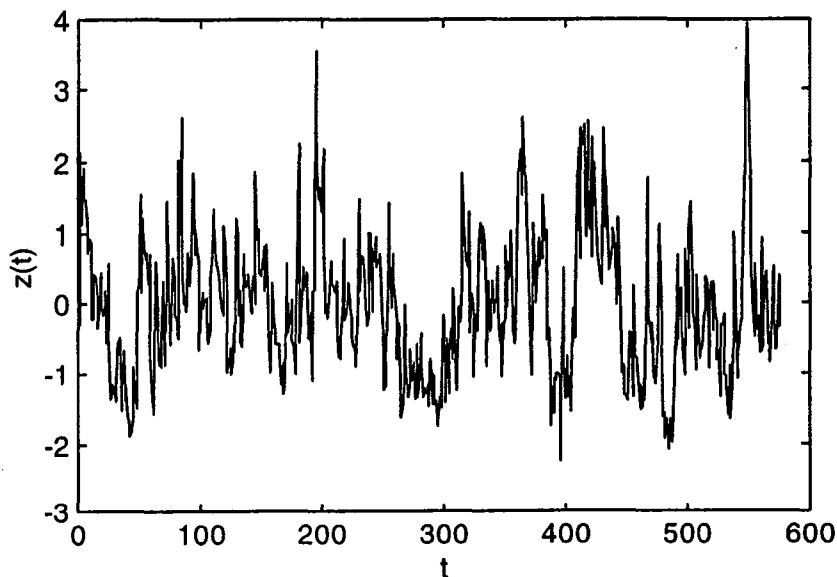


fig. 1 - Gráfico da série de dados reais Furnas com 576 observações.

Tabela 1 - Resultados das estimativas para os parâmetros da série Furnas.

Método	τ	ϕ_1	ϕ_2
Estimat. Mínimos Quadrados	2.20994	0.59602	0.17437
Aproximação das E.M.V.	2.20993	0.59499	0.17376
ARML	2.20993	0.59500	0.17376
Metropolis-Hastings	2.20503	0.59736	0.17588

Vemos que as estimativas dos métodos foram bastante próximas para todos os parâmetros.

Na implementação do algoritmo de Metropolis-Hastings, para cada parâmetro foram simuladas 2 cadeias com 20000 iterações cada e descartada parte inicial de cada cadeia, correspondente a 30% de cada uma. Também foram tomados valores espaçados de 20 iterações (devido à alta correlação entre os valores gerados), totalizando uma amostra de tamanho 1400. Um resumo dos resultados obtidos pelo algoritmo de Metropolis-Hastings são dados na tabela abaixo:

Tabela 2 - Metropolis-Hastings para a série Furnas.

	τ	ϕ_1	ϕ_2
Média	2.20562	0.59322	0.17575
Moda (EMV)	2.20503	0.59736	0.17588
Mediana	2.19634	0.59379	0.17525
$\sqrt{\hat{R}}$ - (crit. de conv.)	1.00462	1.00094	0.99986

Os histogramas construídos com as amostras selecionadas para cada um dos parâmetros, assim como o gráfico de convergência são dados nas figuras seguintes:

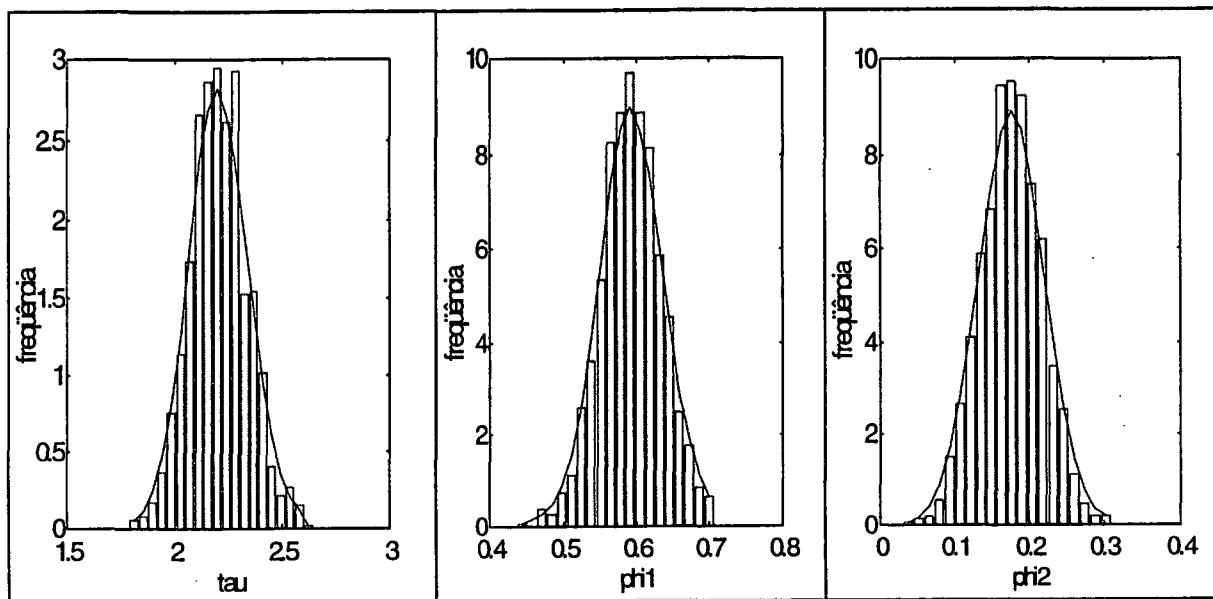


fig. 2 - Histogramas das amostras geradas para τ , ϕ_1 e ϕ_2 .

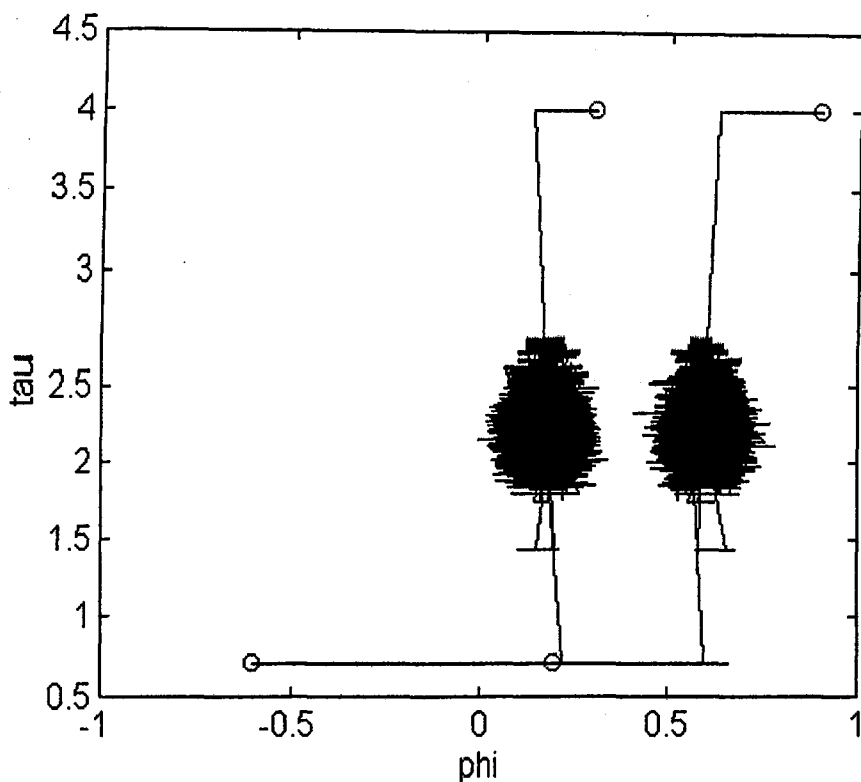


fig. 3 - Gráfico indicando a convergência das 2 cadeias geradas, com seus respectivos pontos de partida.

No gráfico acima, temos no eixo horizontal ϕ_1 e ϕ_2 e no vertical τ . Notamos neste gráfico a convergência do algoritmo para duas condições iniciais diferentes.

Exemplo 2:

Modelo MA(2): série de 100 observações simuladas pelo modelo invertível:
 $z_t = a_t - 0.8a_{t-1} + 0.6a_{t-2}$, onde $a_t \sim N(0, 1)$.

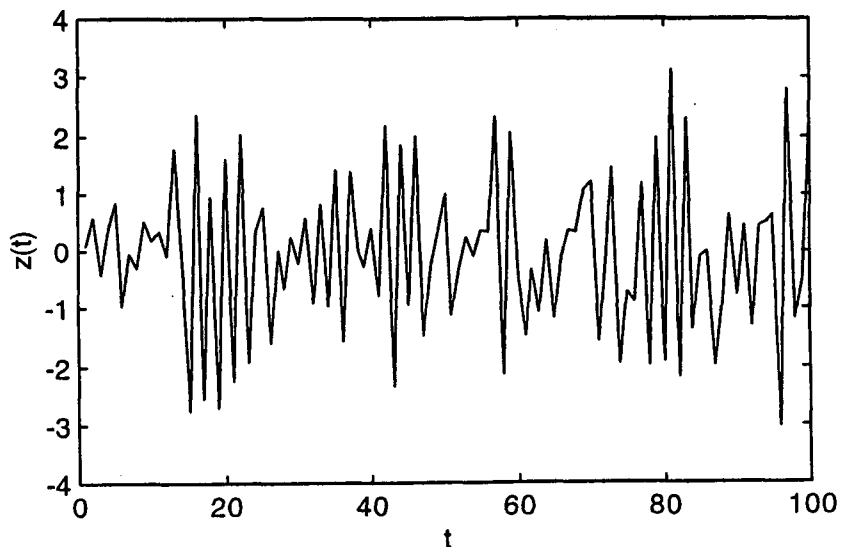


fig. 4 - Gráfico da série simulada $z_t = a_t - 0.8a_{t-1} + 0.6a_{t-2}$.

Para cada um dos métodos, as estimativas calculadas para τ , θ_1 e θ_2 são dadas na tabela abaixo:

Tabela 3 - Resultados das estimativas para os parâmetros da série.

Métodos	τ	θ_1	θ_2
Est. de Máxima Verossimilhança Condicional	0.97237	0.78550	-0.64290
Est. de Máxima Verossimilhança Incondicional	1.00256	0.78630	-0.64340
Gibbs / Metropolis-Hastings	0.99749	0.71545	-0.56344

Analisando os resultados, vemos que para τ as estimativas dos dois últimos métodos foram bastante próximas do valor real $\tau = 1$ e que para θ_1 e θ_2 as estimativas do método MCMC foram mais baixas do que os métodos numéricos.

Para os métodos numéricos, o incremento usado para θ_1 e θ_2 foi de 10^{-4} . Para o método MCMC, foram simuladas 2 cadeias com 15000 iterações cada, para cada um dos parâmetros. A seleção da amostra final foi realizada descartando-se parte inicial de cada cadeia (correspondente a 30% de cada uma) e tomando-se valores espaçados de 10 iterações (devido à correlação entre os valores simulados), totalizando assim 2100 amostras para cada parâmetro. No uso do algoritmo de Metropolis-Hastings para a geração dos candidatos de θ , ajustamos como núcleo uma normal com média $\mu_\theta = (0.7, -0.5)$ e variância $\sigma_\theta^2 = 0.01$. Um resumo dos resultados obtidos pelo algoritmo são dados na tabela abaixo:

Tabela 4 - Gibbs/Metropolis-Hastings para a série $z_t = a_t - 0.8a_{t-1} + 0.6a_{t-2}$.

	τ	θ_1	θ_2
Média	0.98540	0.71043	-0.53684
Moda (EMV)	0.99749	0.71545	-0.56344
Mediana	0.97742	0.71207	-0.53832
$\sqrt{\widehat{R}}$ - (crit. de conv.)	1.00013	1.00911	1.00431

A seguir, apresentamos os histogramas das amostras selecionadas para cada parâmetro e, na figura seguinte, o gráfico de convergência das duas cadeias com todos os pontos simulados pelo algoritmo.

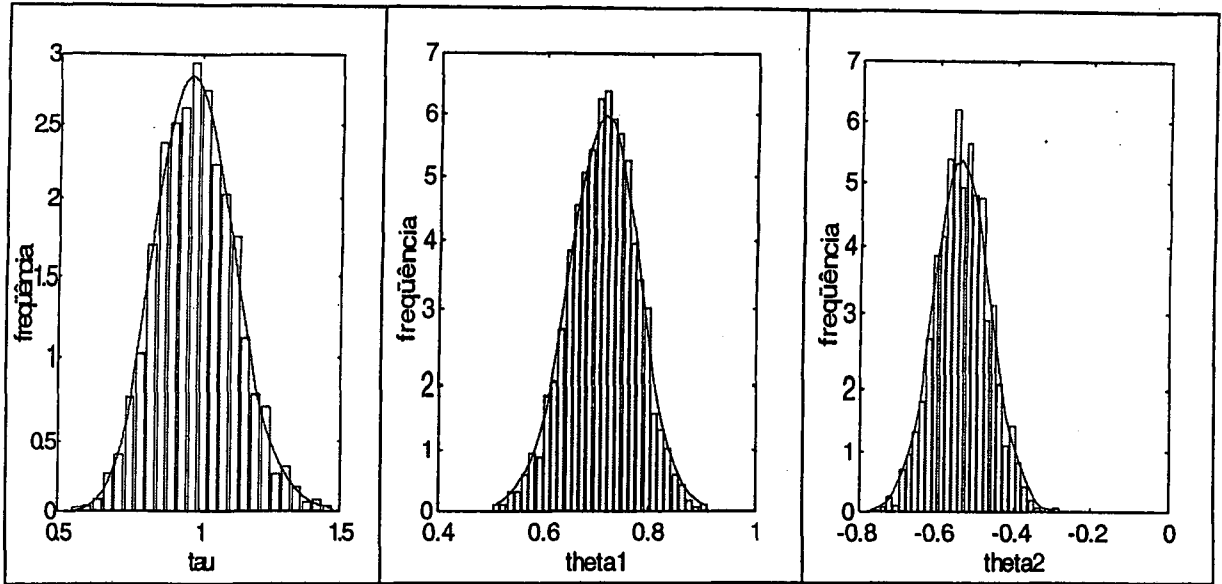


fig. 5 - Histogramas das amostras geradas por MCMC para τ , θ_1 e θ_2 .

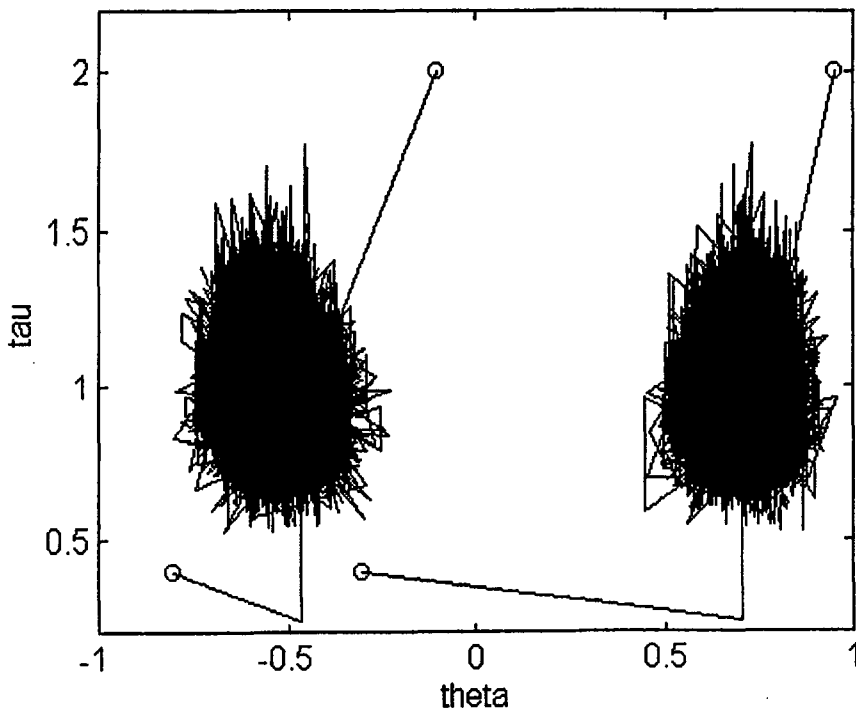


fig. 6 - Gráfico indicando a convergência das duas cadeias geradas, com seus respectivos pontos de partida.

5. Conclusões

As principais conclusões obtidas com o desenvolvimento deste trabalho foram:

- O uso de algoritmos MCMC para o cálculo das estimativas de máxima verossimilhança em modelos AR(p) mostra-se viável quando adota-se a função de

verossimilhança exata, pois tais estimativas só podem ser calculadas fazendo-se simplificações desta função e usando-se métodos numéricos. Com o uso de algoritmos MCMC essas simplificações não são necessárias.

- O uso de algoritmos MCMC em modelos MA(q) para o cálculo das estimativas de máxima verossimilhança apresenta a vantagem de que além de estimar de forma precisa os parâmetros do modelo, não necessita o uso de "backforecasting" como no método estimação de máxima verossimilhança incondicional, o qual é um tanto arbitrário ao se assumir valores nulos para $[e_t]$, $t = n + 1, n + 2, \dots$. Usando MCMC este problema é contornado pois $a_{1-q}, a_{2-q}, \dots, a_0$ são gerados da sua distribuição de probabilidade. Além disso, o uso de MCMC evita um algoritmo de maximização da função $S(\theta)$.

- Na comparação do uso do algoritmo MCMC com os métodos numéricos, os resultados mostraram que as estimativas obtidas via MCMC estão muito próximas das obtidas com os outros métodos no caso dos modelos AR(p). No caso dos modelos MA(q) a diferença entre as estimativas para o parâmetro θ mostraram-se um pouco maiores, porém com resultados bastante satisfatórios. Uma desvantagem que surge no uso de MCMC em modelos MA(q) é o tempo computacional. Devido à necessidade de operações com matrizes de ordem $n \times n$ (grande), estes algoritmos mostram-se mais lentos que os demais. No entanto, a vantagem dos métodos numéricos pode diminuir muito à medida que a discretização do intervalo de variação dos parâmetros nestes métodos aumenta. Esta discretização também afeta a precisão desses métodos.

6. Referências

Anderson, T.W.; Mentz, R.P.; Jarma, N.M.; Martinez, C.I. (1996). *Simulations of Iterative Procedures for Maximum Likelihood Estimation in MA(1) Models*, Commun. Statist. - Simula., 25(4), 851-865.

Box, G.E.; Cox, D.R. (1964). *An Analysis of Transformations*, Journal of the Royal Statistic Society, B, Vol. 6, pp. 211-252.

Box, G.E.; Jenkins, G.M.; Reinsel, G.C. (1994). *Time Series Analysis Forecasting and Control*, 3th ed., Prentice Hall, New Jersey.

Casella, G.; George, E.I. (1992). *Explaining the Gibbs Sampler*, The American Statistician, Vol. 46, No. 3, pp. 167-174.

Chatfield, C. (1989). *The Analysis of Time Series: An Introduction*, 4th ed., Chapman and Hall, New York.

Chib, S.; Greenberg, E. (1995). *Understanding the Metropolis-Hastings Algorithm*, *The American Statistician*, Vol. 49, No. 4, pp. 327-335.

Gelman, A.; Rubin, D.B. (1992). *Inference from Iterative Simulation Using Multiple Sequences*, *Statistical Science*, Vol. 7, No. 4, pp. 457-511.

Geyer, C.J. (1994). *On the Convergence of Monte Carlo Maximum Likelihood Calculations*, *Journal of Royal Statistical Society*, 56, No.1, pp. 261-274.

Geyer, C.J.; Thompson, E.A. (1992). *Constrained Monte Carlo Maximum Likelihood for Dependent Data*, *Journal of Royal Statistical Society*, 54, No. 3, pp. 657-699.

Miller, J.W. (1995). *Exact Maximum Likelihood Estimation in Autoregressive Process*, *Journal of Time Series Analysis*, Vol. 16, No. 6, pp. 607-615.

Morettin, P.A.; Toloi, C.M.C. (1981). *Modelos para Previsão de Séries Temporais*, IMPA, Rio de Janeiro.